5D respiratory motion model based image reconstruction algorithm for 4D cone-beam computed tomography

2015 Inverse Problems 31 115007

(http://iopscience.iop.org/0266-5611/31/11/115007)

IOP Publishing Inverse Problems
Inverse Problems 31 (2015) 115007 (21pp) doi:10.1088/0266-5611/31/11/115007

5D respiratory motion model based image reconstruction algorithm for 4D cone-beam computed tomography

Jiulong Liu[1,2], Xue Zhang[1], Xiaoqun Zhang[1,3], Hongkai Zhao[4], Yu Gao[5], David Thomas[5], Daniel A Low[5] and Hao Gao[1,2]

[1] Department of Mathematics, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China
[2] School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China
[3] Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China
[4] Department of Mathematics, University of California, Irvine, CA 92697, USA
[5] Department of Radiation Oncology, University of California, Los Angeles, CA 90095, USA

E-mail: hao.gao.2012@gmail.com

Received 13 May 2015, revised 26 July 2015
Accepted for publication 26 August 2015
Published 22 October 2015

Abstract

4D cone-beam computed tomography (4DCBCT) reconstructs a temporal sequence of CBCT images for the purpose of motion management or 4D treatment in radiotherapy. However the image reconstruction often involves the binning of projection data to each temporal phase, and therefore suffers from deteriorated image quality due to inaccurate or uneven binning in phase, e.g., under the non-periodic breathing. A 5D model has been developed as an accurate model of (periodic and non-periodic) respiratory motion. That is, given the measurements of breathing amplitude and its time derivative, the 5D model parametrizes the respiratory motion by three time-independent variables, i.e., one reference image and two vector fields. In this work we aim to develop a new 4DCBCT reconstruction method based on 5D model. Instead of reconstructing a temporal sequence of images after the projection binning, the new method reconstructs time-independent reference image and vector fields with no requirement of binning. The image reconstruction is formulated as a optimization problem with total-variation regularization on both reference image and vector fields, and the problem is solved by the proximal alternating minimization algorithm, during which the split Bregman method is used to reconstruct the reference image, and the Chambolle's duality-based algorithm

is used to reconstruct the vector fields. The convergence analysis of the proposed algorithm is provided for this nonconvex problem. Validated by the simulation studies, the new method has significantly improved image reconstruction accuracy due to no binning and reduced number of unknowns via the use of the 5D model.

Keywords: computed tomography, iterative method, image reconstruction

1. Introduction

According to the 2015 report by the American Cancer Society, lung cancer accounts for approximately 27% of all cancer deaths and is by far the leading cause of cancer death among both men and women in the United States. Radiation therapy is commonly used for cancer treatment by conformally delivering the radiation dose to tumors while sparing healthy organs. However, accurate dose planning and delivery remains challenging for lung cancer due to respiratory motion, for which motion management or 4D planning has been extensively studied [1]. In any case, 4D imaging is essential as a prerequisite to guide the dose planning and delivery, such as on-board 4D cone-beam CT (4DCBCT).

4DCBCT provides respiratory phase-resolved volumetric images [2–5]. That is, given the projection data and the breathing amplitude, the projection data is first binned into phases using the breathing measurement, and then each respiratory phase is reconstructed independently. However, due to a very limited amount of projection data available for each phase, its image quality can be degraded which hinders clinical use. To improve image quality by utilizing the prior knowledge that the patient anatomy during respiration is highly correlated, the simultaneous reconstruction methods of all phases are developed with spatiotemporal regularization among different respiratory phases [6–12]; the reconstruction of deformation vector fields on a prior reference image or in addition to the image reconstruction is introduced for 4DCBCT [14–16]. Meanwhile, another approach with no binning requirement, so-called cine CBCT, is developed based on low-rank matrix factorization [17].

On the other hand, a novel breathing motion model, the so-called 5D model, was established to model breathing motion [18] such that the position of a region of interest within the patient can be expressed as a linear function of a reference position vector field, breathing amplitude and its time derivative (rate). That is, given the measurements of breathing amplitude, 5D model parametrizes the respiratory motion by three time-independent variables, i.e., one reference image and two deformation vector fields (corresponding to breathing amplitude and rate). It has been applied to 4D planning CT for motion estimation [13, 19] and image reconstruction [20, 21].

The purpose of this work is to develop a new 4DCBCT reconstruction method based on the 5D model. Instead of reconstructing a temporal sequence of images after the projection binning, the new method reconstructs time-independent reference image and vector fields with no requirement of binning. Compared with the above conventional approaches [2–12, 14–16] that require data binning, the new method does not require binning, and therefore is free from binning artifacts caused by inaccurate or uneven binning in phase, e.g., under non-periodic breathing. Compared with the cine CBCT approach [17] with no data binning requirement as well, the proposed method utilizes the breathing measurement (breathing amplitude and rate), and therefore is expected to improve the image reconstruction quality. Compared with existing approaches with motion reconstruction [14–16], the method here

reconstructs only two time-independent vector fields instead of time-dependent deformation vector fields, and therefore has a significantly reduced number of unknowns.

In terms of reconstruction algorithms, the image reconstruction will be formulated as a nonconvex optimization problem with simultaneous reconstruction of reference image and time-independent vector fields, both of which are regularized by the total variation (TV) [22]. Specifically the proximal alternating minimization [23, 24] will be developed to solve this nonconvex problem, during which the split Bregman method (or so-called alternating direction method of multipliers) is used to reconstruct the reference image [25, 26], and the Chambolle's duality-based algorithm is used to reconstruct the vector fields [27]. The proximal alternating minimization has led to many applications for nonconvex problems [28–30]. Moreover, we will provide the convergence analysis of the proposed algorithm, which is to the best of our knowledge first-of-its-kind for simultaneous reconstruction of image and motion.

The rest of the paper is organized as follows: the background of the proposed method is summarized in section 2; the 5D model based image reconstruction is introduced in section 3; the details of the solution algorithm and its convergence analysis are provided in section 4; the simulation results in comparison with the state-of-art 4DCBCT image reconstruction methods are given in section 5, which demonstrate that the proposed method has significantly improved image quality for periodic or non-periodic breathing; and then the proposed method is summarized in section 6.

2. Background

2.1. 4DCBCT

With the conventional 4DCBCT image reconstruction methods, the projection data is first binned into phases and then the images are reconstructed individually phase by phase [2–5]. Let I_t be the image phases to be reconstructed with T phases of 2D N by N images, i.e., $\{I_t(x_i, y_j),\ 1 \leqslant i, j \leqslant N,\ 1 \leqslant t \leqslant T\}$. The conventional phase-by-phase 4DCBCT method can be formulated as the following iterative reconstruction method with the TV regularization [22]

$$\min_{I_t} \ \| AI_t - y_t \|_2^2 + \mu |\nabla I_t|_1,\ 1 \leqslant t \leqslant T, \tag{1}$$

where A is the system matrix from the x-ray transform [31], y_t the binned projection data to the phase t, and λ the regularization parameter for the spatial TV term (for the purpose of denoising the image during the iterative reconstruction) defined by

$$|\nabla I_t|_1 = \sum_{i,j} \sqrt{\left(\partial_x I_t\right)^2 + \left(\partial_y I_t\right)^2}, \tag{2}$$

with $\partial_x I_t = I_t(x_{i+1}, y_j) - I_t(x_i, y_j)$ and $\partial_y I_t = I_t(x_i, y_{j+1}) - I_t(x_i, y_j)$.

In this work, we also compare with a state-of-art 4DCBCT method that utilizes a priori knowledge that the patient anatomy is correlated for the dynamic images to be reconstructed, through the use of TV to promote the temporal similarity among image phases in addition to the spatial smoothness for each phase [6]. Thus, all the image phases are simultaneously reconstructed instead, which render each other complementary information that would be otherwise missing due to undersampled projection data. That is, the 4DCBCT method with the simultaneous reconstruction of all the image phases is formulated as

$$\min_{\{I_t\}} \sum_t \| AI_t - y_t \|_2^2 + \mu \sum_t |\nabla I_t|_1 + \lambda |\nabla_t I_t|_1, \tag{3}$$

with the temporal TV term defined by

$$|\nabla_t I_t|_1 = \sum_{i,j,t} \left| \partial_t I_t \right| \tag{4}$$

with $\partial_t I_t = I_{t+1}(x_i, y_j) - I_t(x_i, y_j)$ and the regularization parameter λ_t.

For the above convex optimization problems (1) and (3), the solution algorithm can be conveniently developed based on the split Bregman method [25, 26] with the details given in our previous work [8, 9].

2.2. 5D respiratory motion model

The 5D model [18] is developed to accurately model the respiratory motion. It parametrizes the dynamic position of a region of interest on the reference image based on reference position, breathing amplitude and rate. Mathematically, the 5D model can be described by

$$\vec{X}_t = \vec{X} + v_t \vec{M}_1 + f_t \vec{M}_2 \tag{5}$$

in which $\vec{X}_t$ and $\vec{X}$ are spatial coordinates in I_t and the reference image I_0 with the same image intensity value, v_t the breathing amplitude, f_t the breathing rate (v_t and f_t both measured data during free breathing), $\vec{M}_1$ and $\vec{M}_2$ the corresponding time-independent deformation vector fields. Correspondingly

$$I_0\left(\vec{X}\right) = I_t\left(\vec{X}_t\right). \tag{6}$$

That is, the region of interest at $\vec{X}$ in the reference image I_0 deforms to a new location $\vec{X}_t$ in an arbitrary image phase I_t through the 5D model (5) parameterized by the reference image I_0 and its corresponding time-independent vector fields $\vec{M} = (\vec{M}_1, \vec{M}_2)$. Note that $\vec{M}$ is dependent on I_0 and therefore $\vec{M}$ has different values for a different I_0.

2.3. Proximal alternating minimization algorithm

The proximal alternating minimization algorithm [23] has been developed for the nonconvex and nonsmooth minimization problems. Considering a class of nonconvex and nonsmooth problems of the form

$$P(x, y) = f(x) + Q(x, y) + g(y), \tag{7}$$

where

$$\begin{cases} f : R^n \to R \cup \{+\infty\},\ g : R^m \to R \cup \{+\infty\} \text{ are proper lower semicontinous;} \\ Q : R^n \times R^m \to R \text{ is a } C^1 \text{ function;} \\ \nabla Q \text{ is Lipschitz continous on bounded subsets of } R^n \times R^m. \end{cases}$$

The proximal alternating linearized minimization algorithm [23, 24] has been proposed to solve (7), i.e.,

$$\begin{cases} x^{k+1} = \arg\min\limits_x P\left(x, y^k\right) + \frac{1}{2\sigma^k} \| x - x^k \|^2 \\ y^{k+1} = \arg\min\limits_y P\left(x^{k+1}, y\right) + \frac{1}{2\eta^k} \| y - y^k \|^2, \end{cases} \tag{8}$$

where the σ^k and η^k are positive sequences for weighting the proximal term. The convergence of this algorithm can be established based on the Kurdyka–Łojasiewicz (K–Ł) property [23] of the function $P(x, y)$, which will be provided in this work.

3. 5D Model based 4DCBCT

The innovation of this work is to consider the 4DCBCT image reconstruction based on 5D model (5) through

$$\min_{I_0,M} \sum_t \| AI_t\big(I_0, M\big) - y_t \|_2^2 + \mu|\nabla I_0 |_1 + \lambda|\nabla M |_1, \tag{9}$$

where M denotes time-independent vector fields, e.g., $M = (M_{1x}, M_{1y}, M_{2x}, M_{2y})$ in 2D with $\vec{M}_1 = (M_{1x}, M_{1y})$, and $\vec{M}_2 = (M_{2x}, M_{2y})$ correspondingly in (5), and M is regularized in the TV norm component-wise for improved smoothness, i.e., $|\nabla M |_1 = \sum_{i\in\{1,2\},j\in\{x,y\}} |\nabla M_{ij} |_1$. Note that (1) the regularization on M is essential to resolve the illposedness when solving M [32]; (2) no data binning on y_t is assumed here and therefore the number of image phases I_t to be reconstructed is the same as the number of projections y_t. For notation convenience, we shall eliminate $\sum_t$ from the data fidelity term in (9) from now on.

Next we adapt (9) to the linearized form based on which the proximal alternating linearized minimization algorithm can be conveniently developed. Recall that I_t is related to the reference image I_0 and vector fields M by (5) and (6), i.e.

$$I_0\Big(\vec{X}\Big) = I_t\Big(\vec{X} + L_t(M)\Big). \tag{10}$$

With $L_t(M) = v_t\vec{M}_1 + f_t\vec{M}_2$. Then from the linearization

$$I_0\Big(\vec{X}\Big) \approx I_t\Big(\vec{X}\Big) + \nabla^T I_t \cdot L_t(M), \tag{11}$$

where $\nabla^T I_t \cdot L_t(M) = \partial_x I_t \cdot (v_t M_{1x} + f_t M_{2x}) + \partial_y I_t \cdot (v_t M_{1y} + f_t M_{2y})$, we have

$$I_t = I_0 - \nabla^T I_t \cdot L_t(M) \tag{12}$$

and correspondingly the linearized form of (9)

$$\min_{I_0,M} \| A\Big(I_0 - \nabla^T I_t \cdot L_t(M)\Big) - y_t \|_2^2 + \mu|\nabla I_0 |_1 + \lambda|\nabla M |_1. \tag{13}$$

Note that since the vector field M is with respect to I_0, I_t can be computed based on the deformation from I_0, but not vice versa. Also, during the computation of I_t, since the deformed spatial coordinates $\vec{X} + L_t(M)$ from the Cartesian coordinate $\vec{X}$ of I_0 may not be on the Cartesian coordinate of I_t, the interpolation is necessary to derive the Cartesian values of I_t, which is through cubic interpolation here.

To summarize, (9) is solved through iteratively alternating between linearized optimization (13) and aforementioned interpolation, i.e.

$$\begin{cases} \Big(I_0^{K+1}, M^{K+1}\Big) = \arg\min\limits_{I_0,M} \| A\Big(I_0 - \nabla^T I_t^K \cdot L_t(M)\Big) - y_t \|_2^2 \\ \qquad\qquad + \mu|\nabla I_0 |_1 + \lambda|\nabla M |_1 \\ I_t^{K+1} = I_t\Big(I_0^{K+1}, M^{K+1}\Big). \end{cases} \tag{14}$$

Furthermore, we consider the following constrained version of (13)

$$\min_{0\leqslant I_0\leqslant\alpha,-\beta\leqslant M\leqslant\beta} \| A\Big(I_0 - \nabla^T I_t \cdot L_t(M)\Big) - y_t \|_2^2 + \mu|\nabla I_0 |_1 + \lambda|\nabla M |_1, \tag{15}$$

where α and β are positive upper bounds. Note that this constrained formulation (15) is mainly for the convenience of the algorithm convergence analysis. From our simulation experiences, these constraints are not essential for the algorithm to converge in practice.

However this does not exclude the possibility they may be required in the presence of significantly large motions.

4. Alternating reference-image and vector-field reconstruction

The optimization problem (15) is nonconvex and nonsmooth, for which we shall develop the solution algorithm based on the proximal alternating linearized minimization in this section. For the convergence analysis, we also consider the following variant of (15),

$$\min_{0\leqslant I_0\leqslant\alpha,-\beta\leqslant M\leqslant\beta} \| A\big(I_0 - \nabla^T I_0 \cdot L_t(M)\big) - y_t \|_2^2 + \mu|\nabla I_0|_1 + \lambda|\nabla M|_1, \tag{16}$$

or in the iterative form similar to (14)

$$\big(I_0^{K+1}, M^{K+1}\big) = \arg\min_{I_0,M} \| A\big(I_0 - \nabla^T I_0^K \cdot L_t(M)\big) - y_t \|_2^2 + \mu|\nabla I_0|_1 + \lambda|\nabla M|_1, \tag{17}$$

where $\nabla I_t \approx \nabla I_0$ by taking ∇ on both sides of (12) and neglecting the high-order gradients. From our experience, two forms (15) and (16) provide the similar accuracy with (15) being slightly better. In the following, the description of the solution algorithm will be based on (15), while the convergence analysis will be based on (16).

In the notion of general framework (7) for the proximal alternating minimization algorithm, we denote the nonconvex objective function in (15) by

$$P(I_0, M) = f(I_0) + Q(I_0, M) + g(M), \tag{18}$$

where

$$\begin{cases} f(I_0) & = \mu|\nabla I_0|_1 + \mathcal{C}_{0\leqslant I_0\leqslant\alpha}(I_0) \\ Q(I_0, M) & = \| A\big(I_0 - \nabla^T I_t \cdot L_t(M)\big) - y_t \|_2^2, \\ g(M) & = \lambda|\nabla M|_1 + \mathcal{C}_{-\beta\leqslant M\leqslant\beta}(M) \end{cases} \tag{19}$$

with the pointwise indicator function $\mathcal{C}_S$ defined as

$$\mathcal{C}_S(X) = \begin{cases} 0, & \text{if } X(i,j) \in S; \\ +\infty, & \text{if } X(i,j) \notin S. \end{cases} \tag{20}$$

and then apply the proximal alternating linearized minimization (8) for (18), i.e.

$$\begin{cases} I_0^{K+1} = & \arg\min_{0\leqslant I_0\leqslant\alpha} \| A\big(I_0 - \nabla^T I_t^K \cdot L_t\big(M^K\big)\big) - y_t \|_2^2 + \mu|\nabla I_0|_1 \\ & + \dfrac{1}{2\sigma} \| I_0 - I_0^K \|_2^2 \\ M^{K+1} = & \arg\min_{-\beta\leqslant M\leqslant\beta} \| A\big(I_0^{K+1} - \nabla^T I_t^{K+1} \cdot L_t(M)\big) - y_t \|_2^2 \\ & + \lambda|\nabla M|_1 + \frac{1}{2\eta} \| M - M^K \|_2^2. \end{cases} \tag{21}$$

Here K indexes the outer loop by the proximal alternating linearized minimization, and its inner loop (indexed by k) for solving I_0^{K+1} and M^{K+1} will be described next. Note that (21)

also solves (9), since the updates of I_t^{K+1} in (14) are included in (21). Therefore, the inclusion of the iterative scheme (21) in the outer loop (14) makes (21) unchanged. Here we found the constant values for η and σ work well for our problem, although they could vary during iterations (8). Note that although the optimization problem (15) is nonconvex and nonsmooth, the proximal operator added to the subproblems guarantees the convergence of the outer loop [23, 24, 34], which will be proved next.

4.1. Reference-image reconstruction

In this section, we consider the I_0 subproblem of (21), i.e.

$$I_0^{K+1} = \arg\min_{0\leqslant I_0\leqslant\alpha} \| AI_0 - y_t^K \|_2^2 + \mu|\nabla I_0|_1 + \frac{1}{2\sigma}\| I_0 - I_0^K \|_2^2, \tag{22}$$

where $y_t^K = A(\nabla^T I_t^K \cdot L_t(M^K)) + y_t$, and I_t^K is obtained through cubic interpolation on the deformed I_0^K.

Since (22) is a convex problem with sparsity regularization, we adopt the split Bregman method [25, 26] with the details given in our previous work [8, 9]. For algorithm completeness, we shall briefly describe it here. That is, introducing dummy variables $(d_x, d_y) = (\partial_x I_0, \partial_y I_0)$ and its auxiliary variables (b_x, b_y) for isotropic TV norm (2), the convex problem (22) is reformulated as

$$\begin{aligned}\left(I_0, d_x, d_y\right) = \arg\min_{0\leqslant I_0\leqslant\alpha, d_x, d_y} &\| AI_0 - y_t^K \|_2^2 + \mu\left|\left(d_x^k, d_y^k\right)\right|_1 \\ &+ \frac{\rho}{2}\|\left(d_x - \partial_x I_0 - b_x, d_y - \partial_y I_0 - b_y\right)\|_2^2 + \frac{1}{2\sigma}\| I_0 - I_0^K \|_2^2.\end{aligned} \tag{23}$$

Then we update I_0 and (d_x, d_y) alternately through

$$\begin{aligned}I_0^{k+1} = \arg\min_{0\leqslant I_0\leqslant\alpha} &\| AI_0 - y_t^K \|_2^2 + \frac{\rho}{2}\|\left(d_x^k - \partial_x I_0 - b_x^k, d_y^k - \partial_y I_0 - b_y^k\right)\|_2^2 \\ &+ \frac{1}{2\sigma}\| I_0 - I_0^K \|_2^2,\end{aligned} \tag{24}$$

$$\left(d_x^{k+1}, d_y^{k+1}\right) = \arg\min_{d_x, d_y}\left|\left(d_x, d_y\right)\right|_1 + \frac{\rho}{2}\|\left(d_x - \partial_x I_0^k - b_x^k, d_y - \partial_y I_0^k - b_y^k\right)\|_2^2 \tag{25}$$

with auxiliary variables (b_x, b_y) updated by

$$b_x^{k+1} = b_x^k + \left(\partial_x I_0^{k+1} - d_x^{k+1}\right), \tag{26}$$

$$b_y^{k+1} = b_y^k + \left(\partial_y I_0^{k+1} - d_y^{k+1}\right). \tag{27}$$

For the constrained I_0 subproblem (24), we introduce another auxiliary variable D for the constraint to decouple the constraint from the data fidelity, i.e.,

$$\begin{cases} \bar{I}_0^{k+1} &= \arg\min\limits_{I_0} \| AI_0 - y_t^K \|_2^2 + \frac{\rho}{2} \| \left(d_x^k - \partial_x I_0 - b_x^k, d_y^k - \partial_y I_0 - b_y^k \right) \|_2^2 \\ & \quad + \frac{1}{2\sigma} \| I_0 - I_0^K \|_2^2 + \frac{1}{2\gamma} \| I_0 - D^k \|_2^2 \\ I_0^{k+1} &= \arg\min\limits_{0 \leqslant D \leqslant \alpha} \| \bar{I}_0^{k+1} - D \|_2^2. \end{cases} \tag{28}$$

The first equation of (28) is a standard L_2 problem and the second equation of (28) has the analytical solution. Therefore, the solution I_0^{k+1} can be obtained by

$$I_0^{k+1} = \mathbf{\Pi}_{[0,\alpha]} \left(\bar{I}_0^{k+1} \right), \tag{29}$$

where the project operator $\mathbf{\Pi}$ is defined as

$$\mathbf{\Pi}_{[l,u]}(x) = \min \{ \max \{ x, l \}, u \}. \tag{30}$$

Here the linear system for L_2 subproblems in (28) is never explicitly formulated, since the problem (28) can be conveniently solved by conjugate gradient method as described in our previous work [8, 9] without explicitly forming the system matrix A, which can be computed on-the-fly through the parallel computation of x-ray transform and its adjoint [31]. And the (d_x, d_y) subproblem (25) has the explicit solution, i.e., the so-called isotropic shrinkage formula

$$\left(d_x^{k+1}, d_y^{k+1} \right) = \max \left(s^k - \frac{\mu}{\rho}, 0 \right) \frac{\left(\partial_x I_0^k, \partial_y I_0^k \right) + \left(b_x^k, b_y^k \right)}{s^k}, \tag{31}$$

where $s^k = \sqrt{(\partial_x I_0^k + b_x^k)^2 + (\partial_y I_0^k + b_y^k)^2}$.

4.2. Vector-field reconstruction

Next we consider the M subproblem of (21), i.e.

$$M^{K+1} = \arg\min_{-\beta \leqslant M \leqslant \beta} \| A \left(I_0^{K+1} - \nabla^T I_t^K \cdot L_t(M) \right) - y_t \|_2^2 + \lambda |\nabla M|_1 + \frac{1}{2\eta} \| M - M^K \|_2^2 \tag{32}$$

via the Chambolle's duality-based algorithm.

First we introduce a convex relaxation of (32)

$$\begin{aligned} \left(M^{K+1}, U^{K+1} \right) = \arg\min_{-\beta \leqslant M \leqslant \beta, U} & \| A \left(I_0^{(K+1)} - \nabla^T I_t^K \cdot L_t(M) \right) - y_t \|_2^2 + \lambda |\nabla U|_1 \\ & + \frac{1}{2\eta} \| M - M^{(K)} \|_2^2 + \frac{1}{2\theta} \| U - M \|_2^2, \end{aligned} \tag{33}$$

where U is an auxiliary variable. Since (32) is convex, (33) is still convex. Thus, the following alternating scheme (34) is convergent.

$$\begin{cases} U^{k+1} &= \arg\min\limits_{U} \lambda|\nabla U|_1 + \frac{1}{2\theta}\| U - M^k \|_2^2 \\ M^{k+1} &= \arg\min\limits_{-\beta\leqslant M\leqslant\beta} \| A\left(I_0^{K+1} - \nabla^T I_t^K \cdot L_t(M)\right) - y_t \|_2^2 + \frac{1}{2\eta}\| M - M^K \|_2^2 \\ & \quad + \frac{1}{2\theta}\| M - U^{k+1} \|_2^2. \end{cases} \tag{34}$$

Then from the Chambolle's duality-based algorithm [27], the problem (34) can be solved by

$$\begin{cases} U^{k+1} &= \arg\min\limits_{U}|\nabla U|_1 + \frac{1}{2\theta\lambda}\| U - M^k \|_2^2 \\ &= \arg\min\limits_{U}\max\limits_{\|p\|_*\leqslant 1} \langle p, \nabla U\rangle + \frac{1}{2\theta\lambda}\| U - M^k \|_2^2 \\ &= \arg\min\limits_{U}\max\limits_{\|p\|_*\leqslant 1} \langle \mathbf{div}\, p, U^k\rangle + \frac{1}{2\theta\lambda}\| U - M^k \|_2^2 \\ M^{k+1} &= \arg\min\limits_{-\beta\leqslant M\leqslant\beta} \| A\left(I_0^{K+1} - \nabla^T I_t^K \cdot L_t(M)\right) - y_t \|_2^2 + \frac{1}{2\theta}\| M - U^{k+1} \|_2^2 \\ & \quad + \frac{1}{2\eta}\| M - M^K \|_2^2. \end{cases} \tag{35}$$

For the U subproblem, the solution can be analytically obtained by first computing the fixed point of the following iteration (36) over the dual variable p [27, 33]

$$p^{k+1} = \frac{p^k + \tau/\theta\nabla\left(M^{k+1} + \theta\ \mathbf{div}\left(p^k\right)\right)}{1 + \tau/\theta|\ \nabla\left(M^{k+1} + \theta\ \mathbf{div}\left(p^k\right)\right)|} \tag{36}$$

and then

$$U^{k+1} = M^{k+1} + \theta\ \mathbf{div}\left(p^k\right). \tag{37}$$

Note that both (36) and (37) are with respect to each component of M, e.g., $U_{0x}^{k+1} = M_{0x}^{k+1} + \theta\ \mathbf{div}(p_{0x}^k)$ in (37) for M_{0x}.

For the constrained M subproblem, similar to (28), we have

$$\begin{cases} \bar{M}^{k+1} &= \arg\min\limits_{M} \| A\left(I_0^{K+1} - \nabla^T I_t^K \cdot L_t(M)\right) - y_t \|_2^2 + \frac{1}{2\theta}\| M - U^{k+1} \|_2^2 \\ & \quad + \frac{1}{2\eta}\| M - M^K \|_2^2 + \frac{1}{2\gamma}\| M - D^k \|_2^2 \\ M^{k+1} &= \arg\min\limits_{-\beta\leqslant D\leqslant\beta} \| \bar{M}^{k+1} - D \|_2^2, \end{cases} \tag{38}$$

and thus

$$M^{k+1} = \mathbf{\Pi}_{[-\beta,\beta]}\left(\bar{M}^{k+1}\right). \tag{39}$$

The solution $\bar{M}^{k+1}$ for the first L_2 subproblem (38) can be obtained again by conjugate gradient method based on the following first-order optimality conditions

$$\begin{aligned} & - v_t\partial_x I_t \cdot A^T\left(A\left(I_0^K - \nabla^T I_t^K \cdot L_t(M)\right) - y_t\right) + \frac{1}{\theta}\left(M_{1x} - U_{1x}^{k+1}\right) \\ & \quad + \frac{1}{\eta}\left(M_{1x} - M_{1x}^{(K)}\right) + \frac{1}{\nu}\left(M_{1x} - M_{1x}^{(k)}\right) = 0, \end{aligned} \tag{40}$$

$$- v_t \partial_y I_t \cdot A^T \Big(A\Big(I_0^K - \nabla^T I_t^K \cdot L_t(M)\Big) - y_t\Big) + \frac{1}{\theta}\Big(M_{1y} - U_{1y}^{k+1}\Big) + \frac{1}{\eta}\Big(M_{1y} - M_{1y}^{(K)}\Big) + \frac{1}{\nu}\Big(M_{1y} - M_{1y}^{(k)}\Big) = 0, \tag{41}$$

$$- f_t \partial_x I_t \cdot A^T \Big(A\Big(I_0^K - \nabla^T I_t^K \cdot L_t(M)\Big) - y_t\Big) + \frac{1}{\theta}\Big(M_{2x} - U_{2x}^{k+1}\Big) + \frac{1}{\eta}\Big(M_{2x} - M_{2x}^{(K)}\Big) + \frac{1}{\nu}\Big(M_{2x} - M_{2x}^{(k)}\Big) = 0, \tag{42}$$

$$- f_t \partial_y I_t \cdot A^T \Big(A\Big(I_0^K - \nabla^T I_t^K \cdot L_t(M)\Big) - y_t\Big) + \frac{1}{\theta}\Big(M_{2y} - U_{2y}^{k+1}\Big) + \frac{1}{\eta}\Big(M_{2y} - M_{2y}^{(K)}\Big) + \frac{1}{\nu}\Big(M_{2y} - M_{2y}^{(k)}\Big) = 0. \tag{43}$$

Again there is no need to explicitly form the system matrix A, which can be computed on-the-fly through the parallel computation of x-ray transform and its adjoint [31].

4.3. Alternating reconstruction algorithm

For the convenience of implementation, our alternating reference-image and vector-field reconstruction algorithm is summarized in this section.

Algorithm 1. Alternating reference-image and vector-field reconstruction algorithm.

$M^0 = I^0 = \mathbf{0}$
while $\| I_0^{K+1} - I_0^K \|_2 > \delta_1$ **do**
 $I_0^0 = I_0^K$, $b_x^0 = b_y^0 = d_x^0 = d_y^0 = \mathbf{0}$
 while $\| I_0^{k+1} - I_0^k \|_2 > \delta_2$ **do**
 $I_0^{k+1} = \arg\min_{0 \leqslant I_0 \leqslant \alpha} \| A(I_0 - \nabla^T I_t^K \cdot L_t(M^K)) - y_t \|_2^2 + \frac{\rho}{2} \|(d_x^k - \partial_x I_0 - b_x^k, d_y^k - \partial_y I_0 - b_y^k)\|_2^2 + \frac{1}{2\sigma} \| I_0 - I_0^K \|_2^2$
 $d_x^{k+1} = \max\left(s^k - \frac{\mu}{\rho}, 0\right)\frac{\partial_x I_0^k + b_x^k}{s^k}$
 $d_y^{k+1} = \max\left(s^k - \frac{\mu}{\rho}, 0\right)\frac{\partial_y I_0^k + b_y^k}{s^k}$
 $b_x^{k+1} = b_x^k + (\partial_x I_0^{k+1} - d_x^{k+1})$
 $b_y^{k+1} = b_y^k + (\partial_y I_0^{k+1} - d_y^{k+1})$
 $k = k + 1$
end while
 $I_0^{K+1} = I_0^k$
 $M^0 = M^K$, $p^0 = \mathbf{0}$
while $\| M^{k+1} - M^k \|_2 > \delta_3$ **do**
 $p^{k+1} = \frac{p^k + \tau/\theta(M^{k+1} + \theta\, \mathbf{div}(p^k))}{1 + \tau/\theta \,|\, M^{k+1} + \theta\, \mathbf{div}(p^k) \,|}$
 $U^{k+1} = M^k + \theta\, \mathbf{div}(p^k)$
 $M^{k+1} = \arg\min_{-\beta \leqslant M \leqslant \beta} \| A(I_0^{K+1} - \nabla^T I_t^K \cdot L_t(M)) - y_t \|_2^2 + \frac{1}{2\theta} \| M - U^{k+1} \|_2^2 + \frac{1}{2\eta} \| M - M^K \|_2^2$
 $k = k+1$
 end while
 $M^{K+1} = M^k$
 $K = K+1$
end while

In this study, based on algorithm 1, the number of the outer loop K is around 20, the number of split Bregman inner loop for updating I_0 is around 20, the number of Chambolle's inner loop for updating M is around 100, and the number of conjugate gradient iterations in both inner loops is around 10. Similar to other iterative reconstruction methods, the major computational time is spent on x-ray transform and its adjoint. However, the computational time is significantly longer due to the outer loop, which is simply 1 for most iterative methods with image reconstruction only, such as (1) and (3). The computational efficiency will be further explored in the future.

4.4. Convergence analysis

In this section, we shall analyze the convergence of the proposed algorithm for the following optimization problem for which we denote

$$\begin{aligned} P\left(I_0, M\right) = &\parallel A\left(I_0 - \nabla I_0 \cdot L_t(M)\right) - y_t \parallel_2^2 + \mu|\nabla I_0|_1 + \lambda|\nabla M|_1 \\ &+ \mathcal{C}_{0\leqslant I_0\leqslant\alpha}\left(I_0\right) + \mathcal{C}_{-\beta\leqslant M\leqslant\beta}(M) \end{aligned} \tag{44}$$

and

$$\begin{cases} f\left(I_0\right) &= \mu|\nabla I_0|_1 + \mathcal{C}_{0\leqslant I_0\leqslant\alpha}\left(I_0\right) \\ Q\left(I_0, M\right) &= \parallel A\left(I_0 - \nabla^T I_0 \cdot L_t(M)\right) - y_t \parallel_2^2, \\ g(M) &= \lambda|\nabla M|_1 + \mathcal{C}_{-\beta\leqslant M\leqslant\beta}(M). \end{cases} \tag{45}$$

We will start with some notations and preliminary results and then prove the convergence of the sequence (I_0^K, M^K) generated by algorithm 1.

Definition 1. Let f: $R^n \to R \cup \{\infty\}$ be a proper lower semicontinuous function.

(i) For each $x \in \mathrm{dom}\, f$, the Frechet subdifferential of f at x is defined as

$$\hat{\partial} f(x) = \left\{ s\text{: } \liminf_{y\to x, y\neq x} \frac{1}{\parallel x - y \parallel}\left(f(y) - f(x) - \langle s, y - x\rangle\right) \geqslant 0 \right\}. \tag{46}$$

If $x \notin \mathrm{dom}\, f$, then $\hat{\partial} f(x) = \varnothing$.

(ii) The (limiting-)subdifferential of f at $x \in \mathrm{dom}\, f$ is defined as

$$\partial f(x) = \left\{ x^*\text{: } \exists\, x_n \to x, \text{ s.t.} f(x_n) \to f(x),\, s_n \in \hat{\partial} f(x_n) \to x^* \right\}. \tag{47}$$

(iii) The point x is a critical point of f if $0 \in \partial f(x)$.

Definition 2. Let f: $R^n \to R \cup \{\infty\}$ be a proper lower semicontinuous function. f is said to have the K–Ł property at $\bar{x} \in \mathrm{dom}\partial f$ if there exists $\eta > 0$, a neighborhood U of $\bar{x}$ and a continuous concave function ϕ: $[0, \eta) \to R_+$ such that

- $\phi(0) = 0$
- ϕ is C^1 on $(0, \eta)$
- for all $s \in (0, \eta)$, $\phi'(s) > 0$
- for all $x \in U \cap [f(\bar{x} < f(x) < f(\bar{x}) + \eta)]$, the K–Ł inequalityholds

$$\phi'(f(x) - f(\bar{x}))\mathrm{dist}(0, \partial f(x)) > 1.$$

Definition 3. A subset S of $\mathbb{R}^n$ is called the semialgebraic set if there exists a finite number of real polynomial functions g_{ij}, h_{ij} such that

$$S = \bigcup_j \bigcap_i \left\{ x \in \mathbb{R}^n : g_{ij}(x) = 0,\ h_{ij}(x) < 0 \right\}. \tag{48}$$

A function $f(x)$ is called the semi-algebraic function if its graph $\{(x, t) \in \mathbb{R}^n \times \mathbb{R},\ t = f(x)\}$ is a semi-algebraic set.

Theorem 4. *[24] The sequence $Z^K = (x^K, y^K)$ generated by the proximal alternating minimization algorithm for (7) converges to the critical point of (7), if the following conditions hold:*

(1) $P(x, y)$ is a K–Ł function;
(2) $\{Z^K, K = 1, 2, \cdots\}$ is a bounded sequence and there exists some positive constant l, u, such that $l < \sigma^K, \eta^K < u$.
(3) $\nabla Q(x, y)$ has a Lipschitz constant on any bounded set.

Lemma 1. *$P(I_0, M)$ (44) is a K–Ł function.*

Proof 1. Since the indicator function is a real-analytic function [23], $f(I_0) = \mu \left| \nabla I_0 \right|_1 + \mathcal{C}_{0 \leqslant I_0 \leqslant \alpha}(I_0)$ and $g(M) = \lambda \left| \nabla M \right|_1 + \mathcal{C}_{-\beta \leqslant M \leqslant \beta}(M)$ are both real-analytic functions, and therefore, they are both K–Ł functions. In addition, since A and ∇^T are linear operators and L_t is a linear function , $Q(I_0, M) = \| A(I_0^{K+1} - \nabla^T I_0 \cdot L_t(M)) - y_t \|_2^2$ is a real polynomial function, and thus a semi-algebraic function as well. According to [23], the addition of these three terms is a K–Ł function.

Lemma 2. *The sequence $Z^K = (I_0^K, M^K)$ generated by algorithm 1 is a square summable sequence and therefore is subsequence-convergent.*

Proof 2. By definition of proximal operator and the convergence of split Bregman method [25] and Chambolle's duality-based algorithm [27], we have

$$\begin{cases} P\left(I_0^{K+1}, M^K\right) + \frac{1}{2\sigma} \| I_0^{K+1} - I_0^K \|_2^2 \leqslant P\left(I_0^K, M^K\right) \\ P\left(I_0^{K+1}, M^{K+1}\right) + \frac{1}{2\eta} \| M^{K+1} - M^K \|_2^2 \leqslant P\left(I_0^{K+1}, M^K\right) \end{cases} \tag{49}$$

then

$$P\left(Z^K\right) - P\left(Z^{K+1}\right) \geqslant \left(\frac{1}{2\sigma} + \frac{1}{2\eta} \right) \| Z^{i+1} - Z^i \|_2^2. \tag{50}$$

Sum up (50) from 0 to K, and we have

$$P\left(Z^0\right) - P\left(Z^{K+1}\right) \geqslant \left(\frac{1}{2\sigma} + \frac{1}{2\eta} \right) \sum_{i=0}^{i=K} \| Z^{i+1} - Z^i \|_2^2. \tag{51}$$

Let $l = \min\{\sigma, \eta\}$, $u = \max\{\sigma, \eta\}$, and then

$$P\left(Z^0\right) - P\left(Z^{K+1}\right) \geqslant \frac{1}{u} \sum_{i=0}^{i=K} \| Z^{i+1} - Z^i \|_2^2. \tag{52}$$

Since $P(Z^K)$ is a decreasing sequence with lower bound

$$\sum_{i=0}^{i=\infty} \| Z^{i+1} - Z^i \|_2^2 < +\infty. \tag{53}$$

We also have

$$\begin{cases} P\left(I_0^K, M^{K+1}\right) + \frac{1}{2\sigma} \| I_0^{K+1} - I_0^K \|_2^2 \leqslant P\left(I_0, M^K\right) \\ \qquad\qquad + \frac{1}{2\sigma} \| I_0 - I_0^K \|_2^2, \forall\, I_0 \in R^n \\ P\left(I_0^{K+1}, M^{K+1}\right) + \frac{1}{2\eta} \| M^{K+1} - M^K \|_2^2 \leqslant P\left(I_0^K, M\right) \\ \qquad\qquad + \frac{1}{2\eta} \| M - M^K \|_2^2, \forall\, M \in R^m \end{cases} \tag{54}$$

i.e.

$$\begin{cases} f\left(I_0^K\right) + Q\left(I_0^K, M^{K+1}\right) + \frac{1}{2\sigma} \| I_0^{K+1} - I_0^K \|_2^2 \\ \qquad \leqslant f\left(I_0\right) + Q\left(I_0, M^{K+1}\right) + \frac{1}{2\sigma} \| I_0 - I_0^K \|_2^2, \forall\, I_0 \in R^n \\ g\left(M^{K+1}\right) + Q\left(I_0^{K+1}, M^{K+1}\right) + \frac{1}{2\eta} \| M^{K+1} - M^K \|_2^2 \\ \qquad \leqslant g(M) + Q\left(I_0^{K+1}, M\right) + \frac{1}{2\eta} \| M - M^K \|_2^2, \forall\, M \in R^m. \end{cases} \tag{55}$$

Let $(\bar{I}_0, \bar{M})$ be a limit point of sequence (I_0^K, M^K). Then there exists a subsequence $(I_0^{K'}, M^{K'})$ of (I_0^K, M^K) converging to $(\bar{I}_0, \bar{M})$.

$Q(I_0, M)$ is a continuous function, so

$$\lim_{K' \to +\infty} \inf Q\left(I_0^{K'+1}, M^{K'+1}\right) = Q\left(\bar{I}_0, \bar{M}\right). \tag{56}$$

From (55) and (56) together with $\lim_{K \to \infty} \| Z^{K'+1} - Z^{K'} \|_2^2 = 0$, we have

$$\lim_{K' \to +\infty} \inf f\left(I_0^{K'}\right) + Q\left(\bar{I}_0, \bar{M}\right) \leqslant f\left(I_0\right) + Q\left(I_0, \bar{M}\right) + \frac{1}{2\sigma} \| I_0 - \bar{M} \|_2^2, \forall\, I_0 \in R^n. \tag{57}$$

Specifically for $I_0 = \bar{I}_0$, we have

$$\lim_{K' \to +\infty} \inf f\left(I_0^{K'}\right) \leqslant f\left(\bar{I}_0\right). \tag{58}$$

Since $f(I_0) = \mu \left| \nabla^T I_0 \right|_1 + \mathcal{C}_{0 \leqslant I_0 \leqslant \alpha}(I_0)$ is a lower semicontinuous function here, we get

$$\lim_{K' \to +\infty} \inf f\left(I_0^{K'}\right) = f\left(\bar{I}_0\right). \tag{59}$$

Similarly

$$\lim_{K' \to +\infty} \inf g\left(M^{K'}\right) = g\left(\bar{M}\right). \tag{60}$$

Finally

$$\lim_{K' \to +\infty} \inf P\left(I_0^{K'}, M^{K'}\right) = P\left(\bar{I}_0, \bar{M}\right). \tag{61}$$

Therefore, there exists a subsequence of $Z^{K'}$ which is convergent. Furthermore, the existence of l and u is also validated for theorem 4.

Lemma 3. *$\nabla Q(I_0, M)$ in (44) has a Lipschitz constant on any bounded set.*

Proof 3. In the function (44), $Q(I_0, M) = \| A(I_0 - \nabla^T I_0 \cdot L_t(M)) - y_t \|_2^2$, and for simplicity, we denote the $\nabla^T I_0 \cdot L_t(M) \coloneqq \nabla^T I_0 \cdot BM$. It's easy to know B is linear function. Therefore, the gradient of $Q(I_0, M) = \| A(I_0 - \nabla^T I_0 \cdot BM) - y_t \|_2^2$ is

$$\nabla Q(I_0, M) = \begin{bmatrix} \left(A^T - BM \cdot \nabla A^T\right)\left(A\left(I_0 - \nabla^T I_0 \cdot BM\right) - y_t\right) \\ \left(\nabla^T I_0 \cdot B^T A^T\right)\left(A\left(I_0 - \nabla^T I_0 \cdot BM\right) - y_t\right) \end{bmatrix} \tag{62}$$

and its Hessian matrix

$$H = \begin{bmatrix} \left(A^T - BM \cdot \nabla A^T\right)\left(A - A\nabla^T \cdot BM\right) \\ \left(\nabla^T \cdot B^T A^T\right)\left(A\left(I_0 - \nabla^T I_0 \cdot BM\right) - y_t\right) + \left(\nabla^T \cdot B^T A^T\right)\left(A - A\nabla^T I_0 \cdot BM\right) \end{bmatrix}$$

$$\cdots \begin{bmatrix} \left(B \cdot \nabla A^T\right)\left(A\left(I_0 - \nabla^T I_0 \cdot BM\right) - y_t\right) + \left(A^T - BM \cdot \nabla A^T\right)\left(A\nabla^T I_0 \cdot B\right) \\ \left(\nabla^T I_0 \cdot B^T\right)\left(\nabla^T I_0 \cdot B\right) \end{bmatrix}. \tag{63}$$

Since A , ∇ and B are linear bounded operators, we can conclude that the Hessian matrix of $Q(I_0, M)$ (63) are bounded on any bounded set. Namely, $\nabla Q(I_0, M)$ in (44) has a Lipschitz constant on any bounded set.

Lemma 4. *The sequence $Z^K = (I_0^K, M^K)$ generated by algorithm 1 is a bounded sequence.*

Proof 4. Define the bound set $\mathcal{S} = \{0 \leqslant I_0 \leqslant \alpha, -\beta \leqslant M \leqslant \beta\}$. By the definition of project operator $\mathbf{\Pi}$, we conclude that the sequence by algorithm 1 $Z^K = (I_0^K, M_0^K) \in \mathcal{S}$. Therefore, the sequence $Z^K = (I_0^K, M^K)$ generated by algorithm 1 is a bounded sequence.

Theorem 5. *The sequence $Z^K = (I_0^K, M^K)$ generated by algorithm 1 converges to the critical point of $P(I_0, M)$.*

Proof 5. The conclusion follows from theorem 4.

5. Numerical results

The proposed method (abbreviated as '5D Method') (9) was validated based on experimental data of a lung patient in comparison with the conventional phase-by-phase FBP reconstruction (abbreviated as 'FBP') and iterative reconstruction (abbreviated as 'TV') (1) and a state-of-art method of simultaneously reconstructing all phases with spatiotemporal TV regularization (abbreviated as 'TVt') (3).

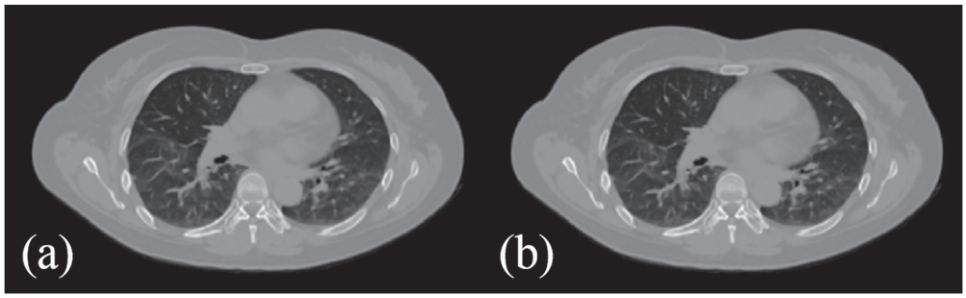

Figure 1. Ground truth. (a) Periodic breathing; (b) non-periodic breathing.

The reconstructed reference image and vector fields, measured breathing amplitude and rate were used to simulate a 4DCBCT in 2D with periodic and non-periodic breathing. Specifically, the breathing amplitude and rate were measured together with a 4DCT patient scan, and then time-independent deformation vector fields of the 5D model were derived based on the existing method [20]. In this proof-of-concept work, the above 5D model and its deformation vector fields were restricted to 2D, i.e., at each projection angle, the 5D model generates a corresponding 2D image, based on which the x-ray transform generates the projection data at this angle, for either periodic or non-periodic data. In our numerical simulation, there were 570 projections evenly distributed in one rotation with 500 detection pixels for each projection. The reconstructed images were 500 × 500, and the displayed images were 500 × 300 (central parts) for better visualization.

For 'FBP', 'TV' and 'TVt', the 570 projections data were first binned into 10 phases. After image reconstruction, the 10 reconstructed image phases were re-mapped to 570 projections in the same order of projection binning. For the proposed '5D Method', no projection binning is necessary and dynamic images I_t were formed based on reconstructed (I_0, M) through the 5D model (5). The reconstruction results were presented with optimized reconstruction parameters.

For quantitative comparison, the reconstruction error was defined by

$$Err = \frac{\| I_{\text{recon}} - I_{\text{Groundtruth}} \|_2}{\| I_{\text{Groundtruth}} \|_2} \tag{65}$$

in L_2 norm. To further illustrate the reconstruction results, the reconstruction errors $|I_{\text{recon}} - I_{\text{Groundtruth}}|$ of all methods with respect to the ground truth (figure 1), and the zoom-in details of the right-bottom images (figure 2 and figure 4) are plotted in figure 3 and figure 5 with the same scale. To summarize, both reconstruction results (figures 2–5) and quantitative errors (tables 1) suggest that the proposed '5D Method' improves the image quality significantly under periodic or non-periodic breathing.

Next we consider the model error to evaluate the robustness of the proposed method, when there is a discrepancy between the ideal 5D model and the realistic model. For this purpose, the projection data is generated from the following perturbed 5D model with 10% relative difference in motion magnitude

$$\vec{X}_t = \vec{X}_0 + (\vec{M}_0 + v_t\vec{M}_1 + f_t\vec{M}_2)\,(1 + 0.1\sin(15\frac{2\pi}{T}t)). \tag{67}$$

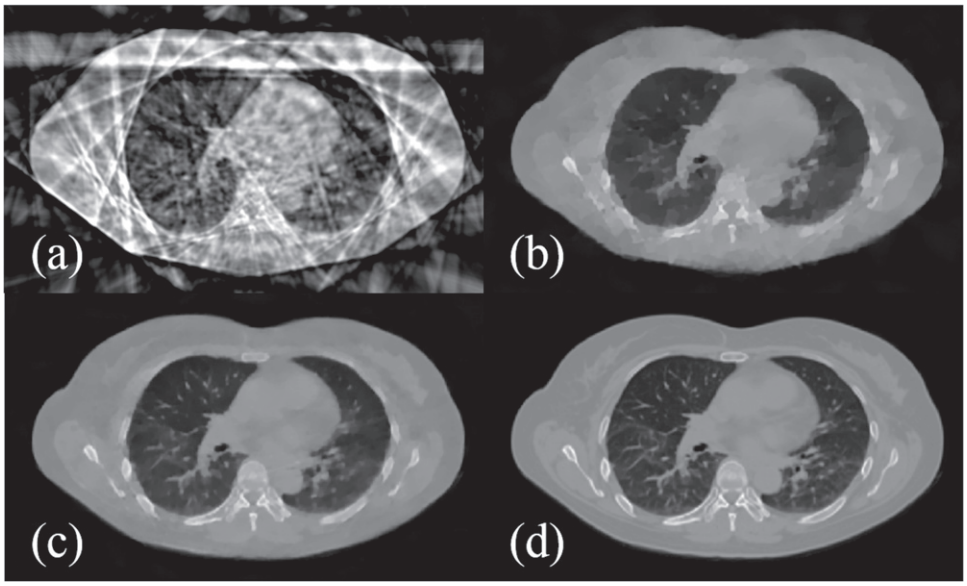

Figure 2. Reconstruction results for periodic breathing. (a) FBP; (b) TV; (c) TVt; (d) 5D Method.

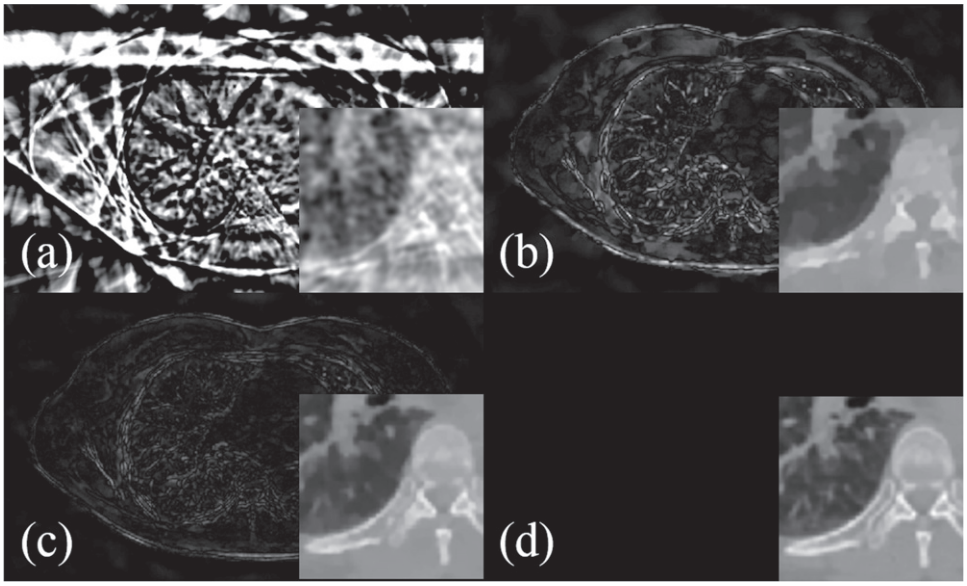

Figure 3. Reconstruction errors and zoom-in details for periodic breathing. (a) FBP; (b) TV; (c) TVt; (d) 5D Method.

Here T is the total number of frames/projections; 15 in (1) is the number of the breathing cycle, assuming 4 seconds per breathing cycle and 60 seconds per scan rotation. Then the proposed method based on the ideal 5D model is used to reconstruct the CBCT images using

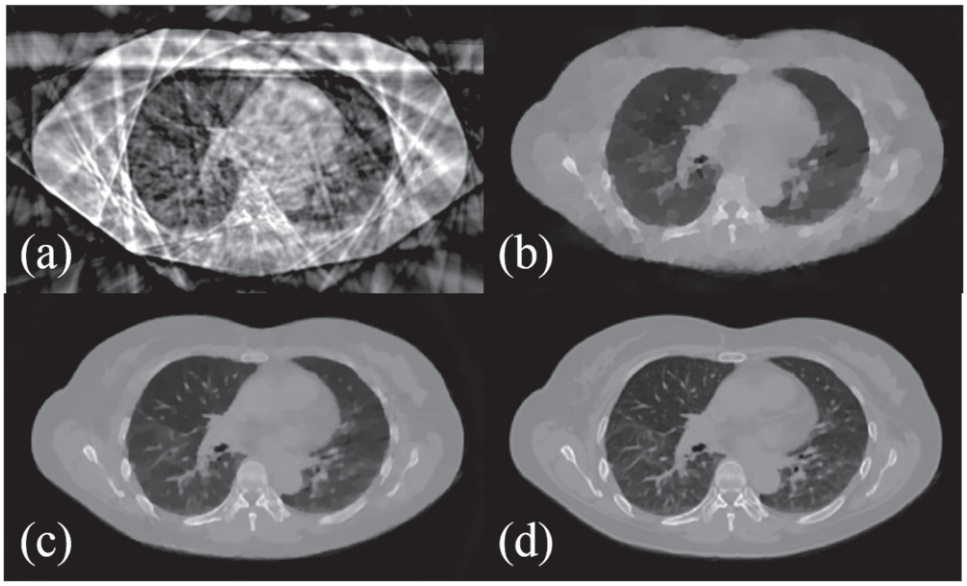

Figure 4. Reconstruction results for non-periodic breathing. (a) FBP; (b) TV; (c) TVt; (d) 5D Method.

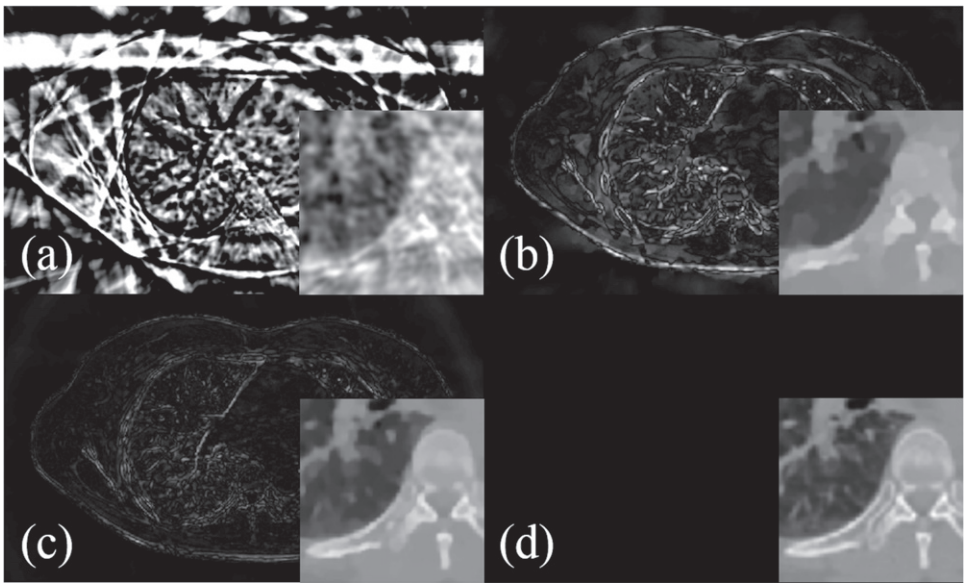

Figure 5. Reconstruction errors and zoom-in details for non-periodic breathing. (a) FBP; (b) TV; (c) TVt; (d) 5D method.

Table 1. Relative errors between reconstructed images and ground truth (periodic breathing) (unit in %).

Method	FBP	TV	TVt	5D method
Periodic	27.63	3.75	2.91	0.44
Non-periodic	28.14	3.86	2.96	0.45

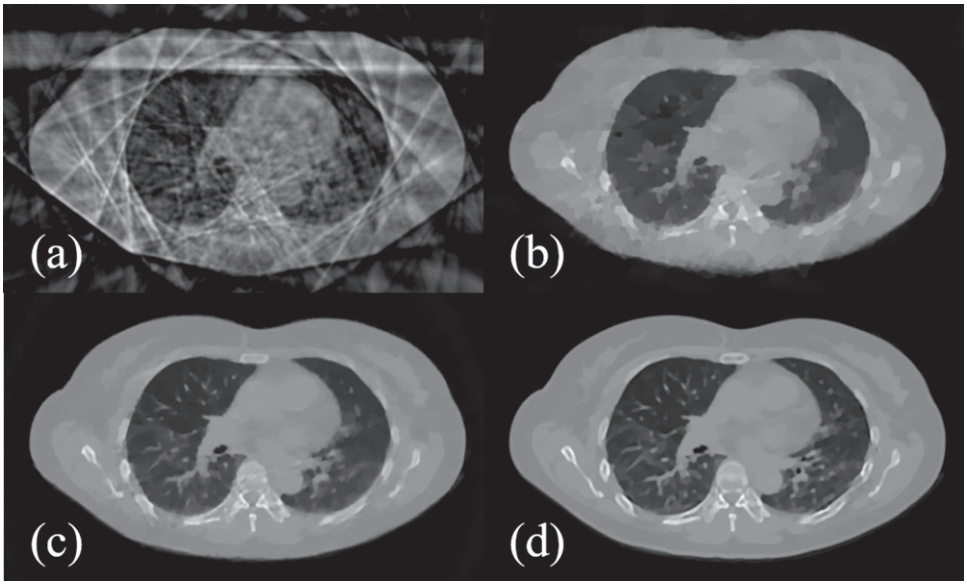

Figure 6. Reconstruction results for periodic breathing (with model error). (a) FBP; (b) TV; (c) TVt; (d) 5D Method.

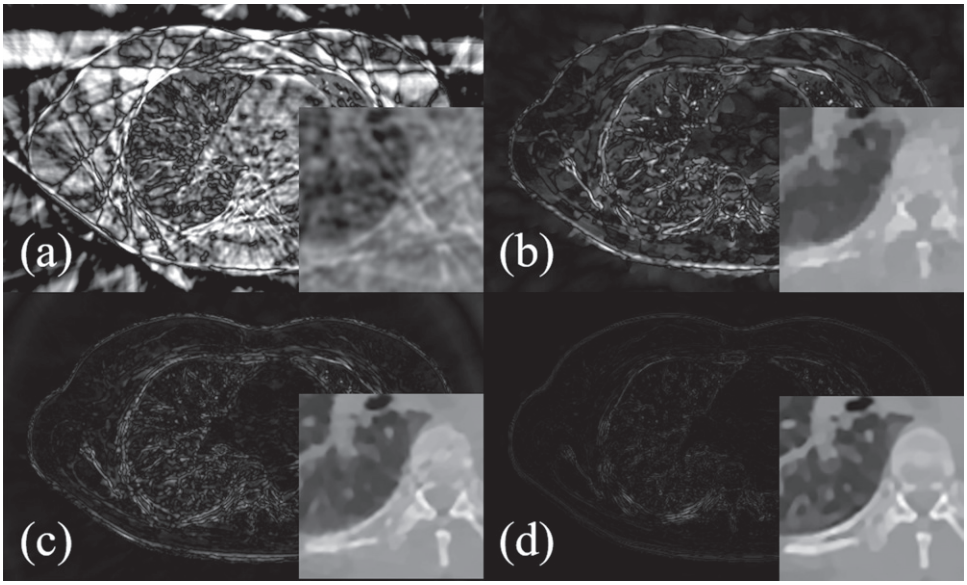

Figure 7. Reconstruction errors and zoom-in details for periodic breathing (with model error). (a) FBP; (b) TV; (c) TVt; (d) 5D Method.

perturbed projection data. The corresponding results are plotted in figures 6, 7, 8, 9, and quantitative errors are summarized in table 2, which again suggest that the proposed "5D Method" improves the image quality under periodic or non-periodic breathing.

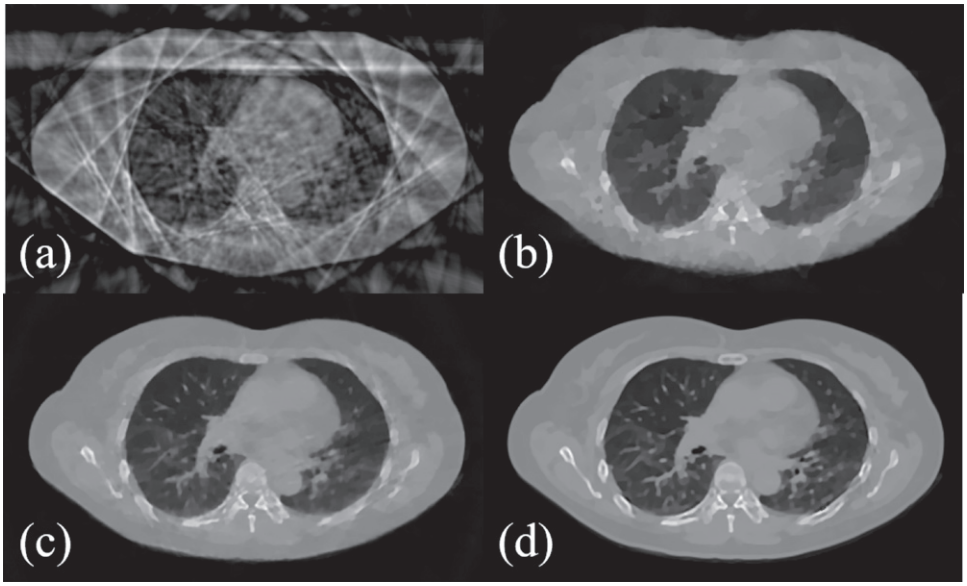

Figure 8. Reconstruction results for non-periodic breathing (with model error). (a) FBP; (b)TV; (c) TVt; (d) 5D Method.

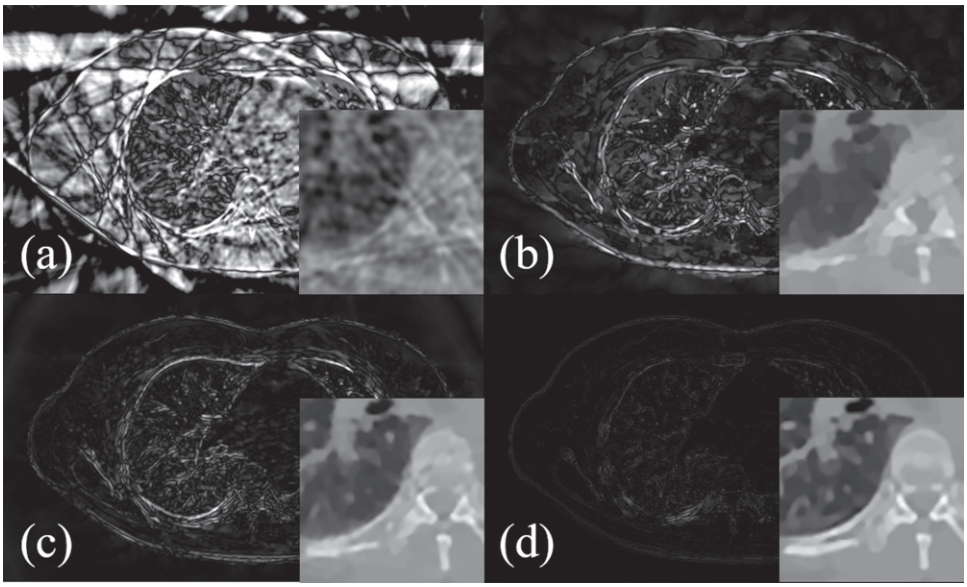

Figure 9. Reconstruction errors and zoom-in details for non-periodic breathing (with model error). (a) FBP; (b) TV; (c) TVt; (d) 5D method.

Table 2. Relative errors between reconstructed images and ground truth (unit in %).

Method	FBP	TV	TVt	5D Method
Periodic	27.82	4.08	2.97	1.33
Non-Periodic	28.04	4.18	3.52	1.33

6. Conclusion

We have developed a new 4DCBCT image reconstruction method based on 5D respiratory motion model, with improved image reconstruction from standard and state-of-art methods for both periodic and non-periodic breathing. The new reconstruction algorithm is formulated as a nonconvex and nonsmooth optimization problem with the reconstruction of reference image and time-independent vector fields, which is solved by the proximal alternating minimization with convergence analysis provided in this work.

Acknowledgments

The authors are grateful to the anonymous reviewers for their valuable comments. Jiulong Liu and Hao Gao were partially supported by the NSFC (#11405105), the 973 Program (#2015CB856000), and the Shanghai Pujiang Talent Program (#14PJ1404500). Daniel A Low was partially supported by the NIH (#R01CA96679).

References

[1] Keall P J, Mageras G S, Balter J M, Emery R S, Forster K M, Jiang S B, Kapatoes J M, Low D A, Murphy M J and Murray B R 2006 The management of respiratory motion in radiation oncology report of AAPM task group 76 *Med. Phys.* **33** 3874–900
[2] Sonke J-J, Zijp L, Remeijer P and van Herk M 2005 Respiratory correlated cone beam CT *Med. Phys.* **32** 1176–86
[3] Kriminski S, Mitschke M, Sorensen S, Wink N M, Chow P E, Tenn S and Solberg T D 2005 Respiratory correlated cone-beam computed tomography on an isocentric C-arm *Phys. Med. Biol.* **50** 5263–80
[4] Dietrich L, Jetter S, Tücking T, Nill S and Oelfke U 2006 Linac-integrated 4D cone beam CT: first experimental results *Phys. Med. Biol.* **51** 2939–52
[5] Li T, Xing L, Munro P, McGuinness C, Chao M, Yang Y, Loo B and Koong A 2006 Four-dimensional cone-beam computed tomography using an on-board imager *Med. Phys.* **33** 3825–33
[6] Chen G-H, Tang J and Leng S 2008 Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets *Med. Phys.* **35** 660–3
[7] Jia X, Lou Y, Dong B, Tian Z and Jiang S 2010 4D computed tomography reconstruction from few-projection data via temporal non-local regularization *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2010* vol 2010 (Berlin: Springer) pp 143–50
[8] Gao H, Cai J-F, Shen Z and Zhao H 2011 Robust principal component analysis-based four-dimensional computed tomography *Phys. Med. Biol.* **56** 3181–98
[9] Gao H, Yu H, Osher S and Wang G 2011 Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM) *Inverse Problems* **27** 115012
[10] Ritschl L, Sawall S, Knaup M, Hess A and Kachelrie M 2012 Iterative 4D cardiac micro-CT image reconstruction using an adaptive spatio-temporal sparsity prior *Phys. Med. Biol.* **57** 1517–25
[11] Wu H, Maier A, Fahrig R and Hornegger J 2012 Spatial-temporal total variation regularization (STTVR) for 4D-CT reconstruction *Int. Soc. Opt. Photonics* **2012** 83133J
[12] Gao H, Li R, Lin Y and Xing L 2012 4D cone beam CT via spatiotemporal tensor framelet *Med. Phys.* **39** 6943–6
[13] Yang D, Lu W, Low D A, Deasy J O, Hope A J and el Naqa I 2008 4D-CT motion estimation using deformable image registration and 5D respiratory motion modeling *Med. Phys.* **35** 4577–90
[14] Christoffersen C P, Hansen D, Poulsen P and Sorensen T S 2013 Registration-based reconstruction of four-dimensional cone beam computed tomography *IEEE Trans. Med. Imaging* **32** 2064–77

[15] Yan H, Zhen X, Folkerts M, Li Y, Pan T, Cervino L, Jiang S B and Jia X 2014 A hybrid reconstruction algorithm for fast and accurate 4D cone-beam CT imaging *Med. Phys.* **41** 071903
[16] Wang J and Gu X 2013 Simultaneous motion estimation and image reconstruction (SMEIR) for 4D cone-beam CT *Med. Phys.* **40** 101912
[17] Cai J, Jia X, Gao H, Jiang S, Shen Z and Zhao H 2014 Cine cone beam CT reconstruction using low-rank matrix factorization: algorithm and a proof-of-principle study *IEEE Trans. Med. Imaging* **33** 1581–91
[18] Low D A, Parikh P J, Lu W, Dempsey J F, Wahab S H, Hubenschmidt J P, Nystrom M M, Handoko M and Bradley J D 2005 Novel breathing motion model for radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **63** 921–9
[19] Zhao T, Lu W, Yang D, Mutic S, Noel C E, Parikh P J, Bradley J D and Low D A 2009 Characterization of free breathing patterns with 5D lung motion model *Med. Phys.* **36** 5183–9
[20] Low D A, White B M, Lee P P, Thomas D H, Gaudio S, Jani S S, Wu X and Lamb J M 2013 A novel CT acquisition and analysis technique for breathing motion modeling *Phys. Med. Biol.* **58** L31
[21] Thomas D, Lamb J, White B, Jani S, Gaudio S, Lee P, Ruan D, McNitt-Gray M and Low D 2014 A novel fast helical 4D-CT acquisition technique to generate low-noise sorting artifact free images at user-selected breathing phases *Int. J. Radiat. Oncol. Biol. Phys.* **89** 191–8
[22] Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica* D **60** 259–68
[23] Bolte J, Sabach S and Teboulle M 2014 Proximal alternating linearized minimization for nonconvex and nonsmooth problems *Math. Program.* **146** 459–94
[24] Attouch H, Bolte J, Redont P and Soubeyran A 2010 Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka–Lojasiewicz inequality *Math. Oper. Res.* **35** 438–57
[25] Goldstein T and Osher S 2009 The split Bregman algorithm for l1 regularized problems *SIAM J. Imaging Sci.* **2** 323–43
[26] Boyd S, Parikh N, Chu E, Peleato B and Eckstein J 2011 Distributed optimization and statistical learning via the alternating direction method of multipliers *Found. Trends Mach. Learn.* **3** 1–122
[27] Chambolle A 2004 An algorithm for total variation minimization and applications *J. Math. Imaging Vis.* **20** 89–97
[28] Xu Y and Yin W 2013 A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion *SIAM J. Imaging Sci.* **6** 1758–89
[29] Ding Q, Zan Y, Huang Q and Zhang X 2015 Dynamic SPECT reconstruction from few projections: a sparsity enforced matrix factorization approach *Inverse Problems* **31** 025004
[30] Bao C, Ji H and Shen Z 2015 Convergence analysis for iterative data-driven tight frame construction scheme *Appl. Comput. Harmon. Anal.* **38** 510–23
[31] Gao H 2012 Fast parallel algorithms for the x-ray transform and its adjoint *Med. Phys.* **39** 7110–20
[32] Horn B K and Schunck B G 1981 Determining optical flow *Technical Symp. East Int. Society for Optics and Photonics* 1981 pp 319–31
[33] Pérez J S, Meinhardt-Llopis E and Facciolo G 2013 TV-L1 optical flow estimation *Image Process. Line* **2013** 137–50
[34] Attouch H and Bolte J 2009 On the convergence of the proximal algorithm for nonsmooth functions involving analytic features *Math. Program.* **116** 5–16